OPERATION OF THE JOINT UCB/USGS DATA CENTER

Award No. 1434-94-G-2387
Barbara Romanowicz

Seismographic Station
University of California, Berkeley
475 McCone Hall
Berkeley, CA 94720-4760

Office: (510) 643-5690, Fax: (510)643-5811, Email: barbara@seismo.berkeley.edu
Program Element: II.2
Key words: Northern California Earthquake Data Center

Annual Project Summary

Investigations Undertaken

The Northern California Earthquake Data Center (NCEDC), a joint project of the Berkeley Seismographic Station and the U.S. Geological Survey at Menlo Park, serves as an "on-line" archive for various types of digital data relating to earthquakes in central and northern California (Table 1). The NCEDC is located at the Seismographic Station, and has been accessible to users via the Internet since mid-1992 (Romanowicz et al., 1994).

The primary goal of the NCEDC is to provide a stable and permanent archival and distribution center of digital waveforms and parametric information for earthquakes in northern and central California. The principal networks contributing seismic data to the data center are the Berkeley Digital Seismic Network (BDSN) operated by the Seismographic Station, and the Northern California Seismic Network (NCSN) short-period network operated by the USGS. The collection of NCSN digital waveforms dates from 1984 to the present, and the BDSN digital data dates from 1987 to the present (Figure 1).

Results

The initial phase of archiving the NCSN earthquake seismograms from 1984 through the present and BDSN data from 1987 through the present is basically complete (Figure 2). All NCSN data tapes have been read and data for all local and regional events has been loaded onto the NCEDC. We are still working with the USGS on the final task of decimating and loading teleseismic events recorded by the NCSN, and reconciling the catalog and parametric information against the list of waveforms to determine what waveforms may still be missing. A list of teleseismic events recorded by the NCSN is now available on the NCEDC. The 16-bit BDSN data from 1987-1991 have been converted to MiniSEED data and is now online.

The archival of current BDSN and NCSN data is an ongoing task. BDSN continuous and event-triggered waveform data is broken into daily archives, run through quality control procedures, and then archived at the NCEDC. NCSN waveform data continue to be collected and processed by the USGS at Menlo Park, and shipped to the NCEDC

on exabyte tapes. Work is underway to provide automatic archiving of NCSN waveforms via a network connection between the NCEDC and the USGS in Menlo Park. Parametric information, such as event catalogs and phase readings from both the BDSN and NCSN networks are automatically updated on the NCEDC on a daily basis.

The loading and redistribution of current NCSN seismograms at the NCEDC experienced two related setbacks this past year. The first problem arose when the NCSN started to incorporate 24-bit seismograms into their seismogram files in June 1995. Until this time, their seismogram files contained only 16-bit data, and were automatically converted from VAX (little-endian) byte order to Sun (big-endian) byte order at the NCEDC under the assumption that all of the seismograms were 16-bit integers. The NCEDC was not notified of this change until several months later, by which time over 250 earthquake files had been erroneously converted by the NCEDC. The NCEDC worked with the USGS and Caltech to define a header format for the seismogram files that would adequately describe the contents and datatypes of the embedded seismograms. The NCSN rewrote both the seismograms and parametric descriptor files for the events in question in the spring of 1996, and continued to use the new format for all events that contain 24-bit seismograms.

The second problem arose when the USGS updated the parametric files for the earthquakes to adopt a uniform station and channel naming convention for the NCSN. The parametric descriptor files for all events that contained 24-bit seismograms were rewritten using the new naming convention. In early July 1996, it was discovered that the station and channel names in the rewritten parametric files were incorrect in ~ 20 percent of the over 191,000 events, so the NCEDC had to revert to using the older parametric files for the NCSN events. Since these older parametric files no longer correspond to the rewritten seismic waveforms, it has prevented the NCSN from importing new events and providing access to seismograms of events later than June 1995. The USGS is currently working to correct the station and channel names in the parametric files, and the NCEDC will resume loading and distributing events as soon as the corrected files are available.

The Seismographic Station has expanded its earthquake monitoring in collaboration with Prof. H. Frank Morrison (Material Science and Mineral Engineering) by monitoring electric and magnetic fields at 2 of UCB's seismic sites (PKD1 and SAO) in Northern California. The 3 channels of magnetic and 2 channels of electric data at each site are telemetered in real-time along with seismic data to the Seismographic Station, where they are processed and archived at the NCEDC in a similar fashion to the seismic data. The 40 Hz data channels will remain online for ~ 6-12 months, while the 1 Hz data will remain online permanently. Work is underway to archive and distribute Schumann resonance parameters derived from this dataset.

The NCEDC is continuing to expand its role through the inclusion of other related geophysical datasets for central and northern California. The Bay Area Regional Deformation (BARD) network of continuously monitored GPS receivers in Northern California now incorporates 22 sites in Northern California. There are currently six sites operated by UC Berkeley (one contributed by LLNL), five sites operated by the USGS, four sites operated by JPL, three sites operated by the US Coast Guard, two sites operated by Trimble Navigation, and 2 sites operated by Scripps Institute of Oceanography. Data from the Berkeley sites are automatically downloaded and archived at the NCEDC on a daily basis. Data from the USGS sites are manually downloaded and transferred to the NCEDC on a daily basis, and are automatically archived by the NCEDC. The other sites are acquired from their respective operators, and are archived by the NCEDC.

The Unocal Corporation operates a micro-seismic monitoring network in the Geysers regions of northern California, and has agreed to release six years of triggered waveform

data (1987-1994) for archival and distribution at the NCEDC. This dataset represents roughly 100,000 events that were recorded by the Unocal Geysers network, and is currently available via research accounts at the NCEDC. Although no parametric information such as hypocenters or phase readings is currently being made available by Unocal, several scientists have already deemed this dataset as a useful addition to the NCEDC. The Seismographic Station, LBNL, and NCEDC are working with Unocal to try to acquire the catalog of earthquake locations and magnitudes.

The NCEDC, in conjunction with the Council of the National Seismic System (CNSS), is producing and distributing a world-wide composite catalog of earthquakes based on the catalogs of the national and various US regional networks. Each network updates their earthquake catalog on a daily basis at the NCEDC, and the NCEDC constructs a composite world-wide earthquake catalog by combining the data, removing duplicate entries that may occur from multiple networks recording an event, and giving priority to the data from each network's "authoritative region". The catalog is available for use at the NCEDC, and is made available to anyone over the Internet.

The NCEDC continues to make significant use of the World Wide Web for accessing and retrieving data. The World Wide Web is comprised of network-accessible information that uses hypertext and multi-media techniques to deliver information and to provide links between different documents and sources of information. World Wide Web browsers provide a consistent and easy-to-use interface that allows users to easily explore, request, and retrieve information that is spread across the entire Internet.

Availability of data

The NCEDC continues to use the World Wide Web as a principal interface for users to request, search, and receive data from the NCEDC. The NCEDC has implemented a number of useful and original mechanisms of data search and retrieval using the World Wide Web, and are available to anyone on the Internet. All of the documentation about the NCEDC, including the research users' guide, is available via the Web. Users can perform catalog searches and retrieve hypocentral information and phase readings from the various earthquake catalogs at the NCEDC via easy-to-use forms on the Web. In addition, users can peruse the index of available broadband data at the NCEDC, and can request and retrieve broadband data in standard SEED format via the Web. Access to all datasets is available via research accounts at the NCEDC. The NCEDC's home page address is "http://quake.geo.berkeley.edu/". Contact person: Mr Douglas Neuhauser (e-mail address: doug@seismo.berkeley.edu).

OPERATION OF THE JOINT UCB/USGS DATA CENTER

Award No. 1434-94-G-2387

Barbara Romanowicz

Seismographic Station
University of California, Berkeley
475 McCone Hall
Berkeley, CA 94720-4760

Office: (510) 643-5690, Fax: (510)643-5811, Email: barbara@seismo.berkeley.edu

Program Element: II.2
Key words: Northern California Earthquake Data Center

Non-technical Project Summary

The Northern California Earthquake Data Center (NCEDC) is serving as on-line archive and distribution facility for waveform and catalog data for several regional networks: The Northern California Seismic Network (NCSN) operated by the U.S. Geological Survey (NCSN), the Berkeley Digital Seismic Network (BDSN) operated by the U.C. Berkeley Seismographic Station, and the Bay Area Deformation Array (BARD) operated jointly by UC Berkeley, USGS with participation of several other San Francisco Bay Area institutions. These data serve as basis for many research projects relevant to NEHRP goals in the Bay Area and more generally central and Northern California.

Table 1

Volume of data archived at NCEDC by data type	
Data Stored	MBytes
Broadband seismograms (compressed)	189,337
NCSN event seismograms (compressed)	156,107
Electric and Magnetic field waveforms (compressed)	17,422
GPS data (compressed and raw)	21,944
Unocal Geysers regions seismograms	17,319
Misc data	15,194
Total size of archived data	417,323

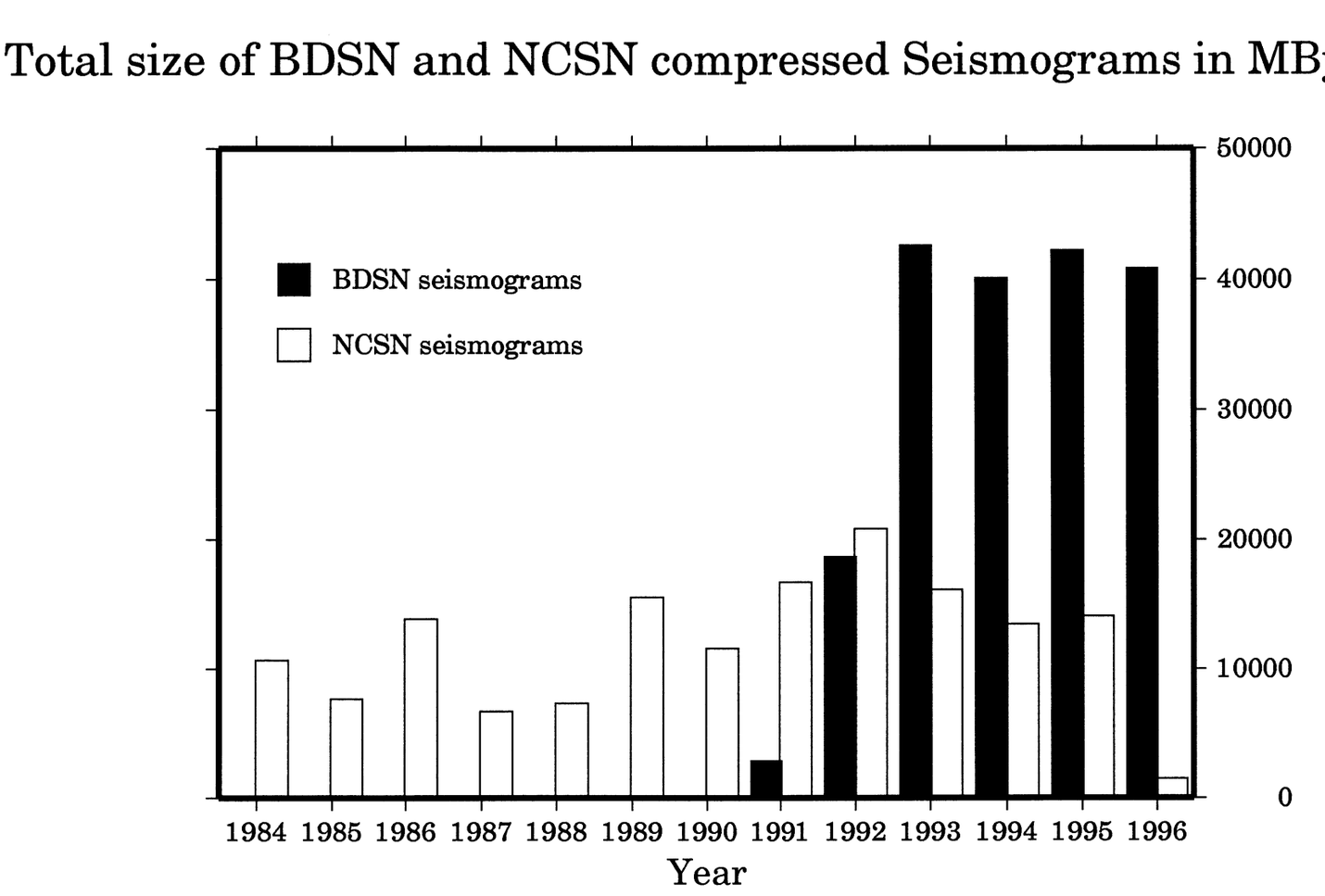

Figure 1: Megabytes of compressed seismograms loaded on NCEDC plotted by year from 1984 to August 1996. *Courtesy of S. Fulton and D. Neuhauser.*

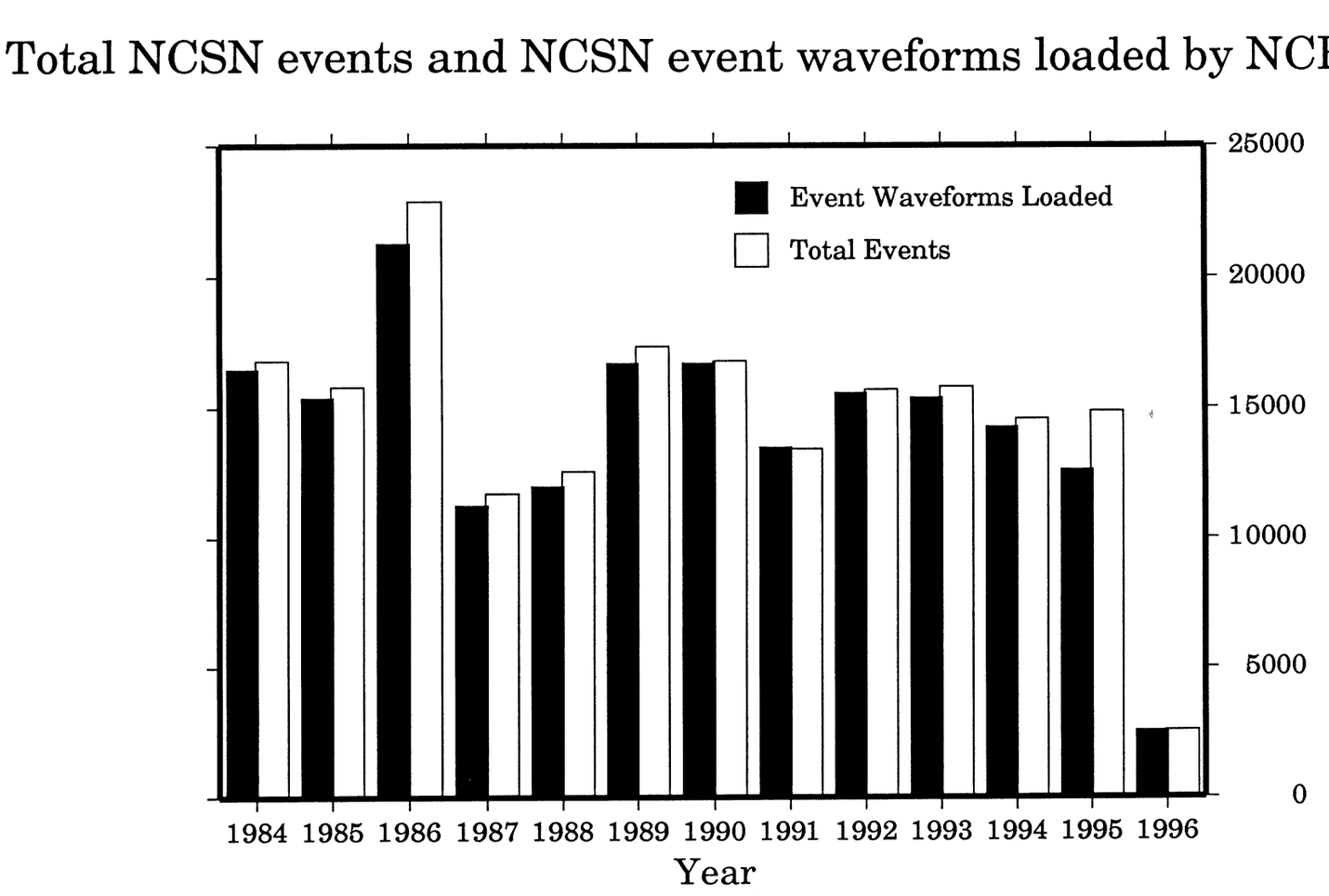

Figure 2: Comparison of total number of events acquired by NCSN for each year to the number of events loaded on the NCEDC through August 1996. Data from prior years must be copied from NCSN archive tapes before being sent to the data center. *Courtesy of S. Fulton and D. Neuhauser.*